The cost of parallel consolidation into visual working memory

Reuben Rideaux Research School of Psychology, The Australian National University, Canberra, Australia

Mark Edwards Research School of Psychology, The Australian National University, Canberra, Australia

A growing body of evidence indicates that information can be consolidated into visual working memory in parallel. Initially, it was suggested that color information could be consolidated in parallel while orientation was strictly limited to serial consolidation (Liu & Becker, 2013). However, we recently found evidence suggesting that both orientation and motion direction items can be consolidated in parallel, with different levels of accuracy (Rideaux, Apthorp, & Edwards, 2015). Here we examine whether there is a cost associated with parallel consolidation of orientation and direction information by comparing performance, in terms of precision and guess rate, on a target recall task where items are presented either sequentially or simultaneously. The results compellingly indicate that motion direction can be consolidated in parallel, but the evidence for orientation is less conclusive. Further, we find that there is a twofold cost associated with parallel consolidation of direction: Both the probability of failing to consolidate one (or both) item/s increases and the precision at which representations are encoded is reduced. Additionally, we find evidence indicating that the increased consolidation failure may be due to interference between items presented simultaneously, and is moderated by item similarity. These findings suggest that a biased competition model may explain differences in parallel consolidation between features.

Introduction

Whereas our environment is visually rich, only a small proportion of the information that enters the retina is stored as a durable representation. Information is initially stored in sensory (iconic) memory, which is characterized as high capacity memory whose contents decay within a few hundred milliseconds (Sperling, 1960, 1963). Following this, a small proportion the information stored in sensory memory is transferred to visual working memory (VWM), aka visual short-term memory (Cowan, 2001, 2010). While determining the precise capacity of VWM, and the nature of this capacity, has been the focus of a vast number of studies (for a review, see Ma, Husain, and Bays, 2014), another important aspect of VWM that has been drawing progressively more attention is the process of consolidation, i.e., the formation of VWM representations.

Initially, research suggested that consolidation of information into VWM was restricted to serial processing, i.e., only one item could be consolidated at a time (Huang, Treisman, & Pashler, 2007). However, recently several studies have found evidence indicating that parallel consolidation of color is possible, albeit restricted to two or three items (Mance, Becker, & Liu, 2012). Whereas initially it was suggested that the capacity to consolidate information in parallel may be limited to color (Becker, Miller, & Liu, 2013; Liu & Becker, 2013), we recently found compelling evidence that it is also possible with motion direction, and some evidence to suggest it may even be possible with orientation, at different levels of accuracy (Rideaux, Apthorp, & Edwards, 2015). One potential explanation we flagged for this difference in accuracy is that the precision of items consolidated in parallel (compared to serially) is reduced.

In our previous study, a matching task was employed in which observers were presented with a number of items and then required to indicate whether a subsequently presented item was among them (present) or not (absent) (Rideaux et al., 2015). We found that performance was the same for two motion direction items presented sequentially or simultaneously. However, when the range of items used was reduced from 360° to 180°, while serial consolidation performance remained the same, parallel consolidation suffered. We then examined performance on the task using orientation and found it was similar to that in the second reduced range motion direction condition:

Citation: Rideaux, R., & Edwards, M. (2016). The cost of parallel consolidation into visual working memory. *Journal of Vision*, *16*(6):1, 1–14, doi:10.1167/16.6.1.

doi: 10.1167/16.6.1 Received November 23, 2015; published April 1, 2016 ISSN 1534-7362

worse in the simultaneous condition; yet better than predicted by purely serial consolidation.

To account for these findings, we proposed that items consolidated in parallel may be encoded at a lower precision than those consolidated serially, and that the difference in performance between sequential versus simultaneous conditions reflects an interaction between this reduction in precision and the similarity of items within a relatively small perceptual space. That is, as the precision of encoded representations relative to the separation of items along a feature dimension is reduced, the probability of them being mistaken for neighboring items (during the decision stage of the task) is increased. In line with this, Umemoto, Drew, Ester, and Awh (2010) found evidence that the precision of multiple (four) orientation items was reduced when presented simultaneously (relative to sequentially). However, the authors were not explicitly controlling for serial consolidation and given the exposure duration (300 ms) and lack of backward masking in their experiment: Observers likely employed a serial strategy even in the simultaneous condition. The notion of a cost associated with processing multiple items is not unusual: in the motion processing literature the cost associated with processing two (or more) motion direction signals simultaneously (relative to one signal) is well established (Edwards & Greenwood, 2005; Edwards & Rideaux, 2013; Qian, Andersen, & Adelson, 1994; Rideaux & Edwards, 2014).

However, another potential source of error which may account for the poorer performance observed in the simultaneous condition on the matching task may have been an increase in consolidation failure. A number of studies have demonstrated that when items are presented simultaneously, as opposed to sequentially, competition between items can result in consolidation failure (Ihssen, Linden, & Shapiro, 2010; Scalf & Beck, 2010). Indeed, such findings have prompted the claim that this competition is directly responsible for the capacity of VWM, i.e., the biased competition model of VWM (Shapiro & Miller, 2011), and are supported by neuroimaging studies which show a reduced BOLD signal when items are presented simultaneously compared to sequentially (Beck & Kastner, 2007; Kastner & Ungerleider, 2001).Originally proposed by Desimone and Duncan (1995) to explain the capacity of visual selective attention, the general principle of the biased competition model is that items within the visual field compete for representation within the limited capacity of regions (aka content maps) in the visual cortex. These regions can be conceptualized as two-dimensional areas of the cortex with coherent spatial organization where the preferred stimuli of neurons change smoothly from one location to the next, e.g., area MT where neurons vary in motion direction selectivity (Albright, Desimone, & Gross, 1984). According to this account, a number of factors moderate the degree of competition between items including the size of the receptive fields in visual areas, the number of items, item similarity, and item spatial proximity (Franconeri, Alvarez, & Cavanagh, 2013; Kastner & Ungerleider, 2001). In contrast to unlimited parallel models that claim no loss of accuracy or increased consolidation failure, e.g., the consolidation bandwidth model (Miller, Becker, & Liu, 2014), this model would predict that the likelihood of consolidation failure may increase when items are presented simultaneously.

In summary, there are two potential sources of error that may account for poorer performance on a matching task when items are presented simultaneously compared to sequentially: The precision of encoded items may be reduced and/or the likelihood of consolidation failure may be increased. The most compelling evidence for parallel consolidation is a reduction in the precision of encoded items. In contrast, strict serial consolidation would predict no change in precision and an inability to encode more than one item on each trial, i.e., a 50% consolidation failure rate. However, a mixture of these resulting from parallel consolidation can also be explained, under a biased competition framework. Here we explicitly examine the sources of error, in terms of precision and consolidation failure, associated with attempting to consolidate motion direction and orientation in parallel. Although we found partial evidence to indicate that orientation can be consolidated in parallel, this conflicts with previous findings (Liu & Becker, 2013); thus, this test will also serve to clarify whether there is a flexible time-accuracy trade-off associated with parallel consolidation, i.e., reduced precision, or if one or both of these features are strictly limited to serial processing, i.e., increased (50%) consolidation failure.

Experiment 1

We previously found evidence suggesting that both orientation and motion direction information can be consolidated into VWM in parallel (Rideaux et al., 2015). Furthermore, the results suggested that items consolidated in parallel are encoded at a lower precision than those consolidated serially, which may account for the differential performance of parallel consolidation observed for different types of information. Here we explicitly examine the source of error associated with attempting to consolidate these features in parallel to determine whether there is a cost to the precision of items encoded in parallel, or if they are limited to serial processing.

Method

Observers

Twenty-four observers participated in the study (mean age, 22). All had normal or corrected-to-normal acuity and gave informed written consent to participate in the study. All observers were compensated $20 for participation.

Apparatus

Experiments were run under the MATLAB (version R2013a) programming environment, using software from the PsychToolbox (Brainard, 1997; Pelli, 1997). Stimuli were presented on a Phillips Brilliance 202P4 CRT monitor that was driven by an Intel Iris graphics card in a host MacBook Pro computer. The monitor had a spatial resolution of 1024×768 pixels and a frame rate of 120 Hz.

Stimuli

The stimuli and procedure were similar to those employed by Liu and Becker (2013). A 2×2 experimental design was used: presentation (sequential/simultaneous) $\times$ feature (orientation/motion direction). The general presentation sequence consisted of a stimulus interval/s followed by a cue, and then a response interval. In each stimulus interval two items were presented either sequentially or simultaneously (40 ms) followed by a mask (200 ms) to prevent further processing from iconic memory. This exposure duration was chosen as it is approximately the average duration required to consolidate a single item, or two items in parallel (Mance et al., 2012; Rideaux et al., 2015). Items presented sequentially were separated by a 500 ms interstimulus interval (ISI) where only the fixation cross was present. Items/masks were presented 8° (visual angle) to the left and right of fixation, with the left item always presented first in the sequential condition.

In the orientation condition the stimuli were Gabors (contrast, 0.7; spatial frequency, 1 cycles/°; random phase) within a Gaussian envelope (4° radius) and the mask was pixel noise of random luminance levels with a uniform distribution (0 – 63 cd/m^2) within a circular aperture (4.2° radius). In the motion direction condition stimuli consisted of 100 Gaussian blobs (0.3° radius), to allow subpixel resolution movement, within a circular aperture (4° radius), which wrapped around when they reached the edge of the aperture. The blobs were displaced 0.082° each frame, resulting in a speed of 9.8°/s. The mask in this condition consisted of 300 Gaussian blobs within a circular aperture (4.2° radius), positioned randomly on each frame, creating a percept of dynamic noise. The orientation/direction of each pair of items was determined pseudorandomly from between 0°–179° and 0°–359°, respectively, with the constraint that they must be separated by at least 15°.

The stimulus interval/s and cue were separated by a 500 ms ISI. The cue consisted of a white circle (500 ms) presented in the location of one of the items. The location was determined pseudorandomly such that the cue appeared in each location an equal number of times. Finally, the cue was followed by a response interval consisting of either a Gabor, identical to those used in the stimulus interval, (orientation condition) or an arrow (length 6°) extending from fixation (direction condition). During the response interval, the cursor became visible and the orientation/direction of the Gabor/arrow could be manipulated by moving the mouse. Examples of the presentation sequence are illustrated in Figure 1.

The background was grey (mean luminance, 12 cd/m^2) and the blobs in the direction condition were white (mean luminance, 63 cd/m^2). The observer sat 50 cm from the monitor, with the head supported on a chin rest.

Procedure

The observer's task was to match the orientation/direction of the Gabor/arrow in the response interval to that of the cued item in the preceding stimulus interval. No duration limit was used to restrict responses: Once the observer had moved the Gabor/arrow with the mouse to the orientation/direction they believed matched that of the cued item, they would left-click to indicate their response and initiate the next trial. Observers were instructed to maintain fixation throughout the presentation sequence and to remember both items in order to perform the task accurately.

Observers were randomly split into two groups; half were run in the orientation condition and the other half in the direction condition. Initially, observers spent approximately 10 min performing the task without recording data, in order to familiarize them with the stimuli/task. Following this, each observer ran six blocks of each presentation condition, i.e., sequential/simultaneous, randomly interleaved within a mega block. Each block consisted of 100 trials, totalling 1200 trials and an approximate testing duration of one hour per observer.

Data analysis

For each trial, the offset (error) was calculated by subtracting the orientation/direction recorded from the observer's response from that of the cued item. Initially, the raw mean and variance of the offset was analyzed for each participant. However, there are two sources of variability within the offsets, resulting from two types of trials: one where the observer successfully

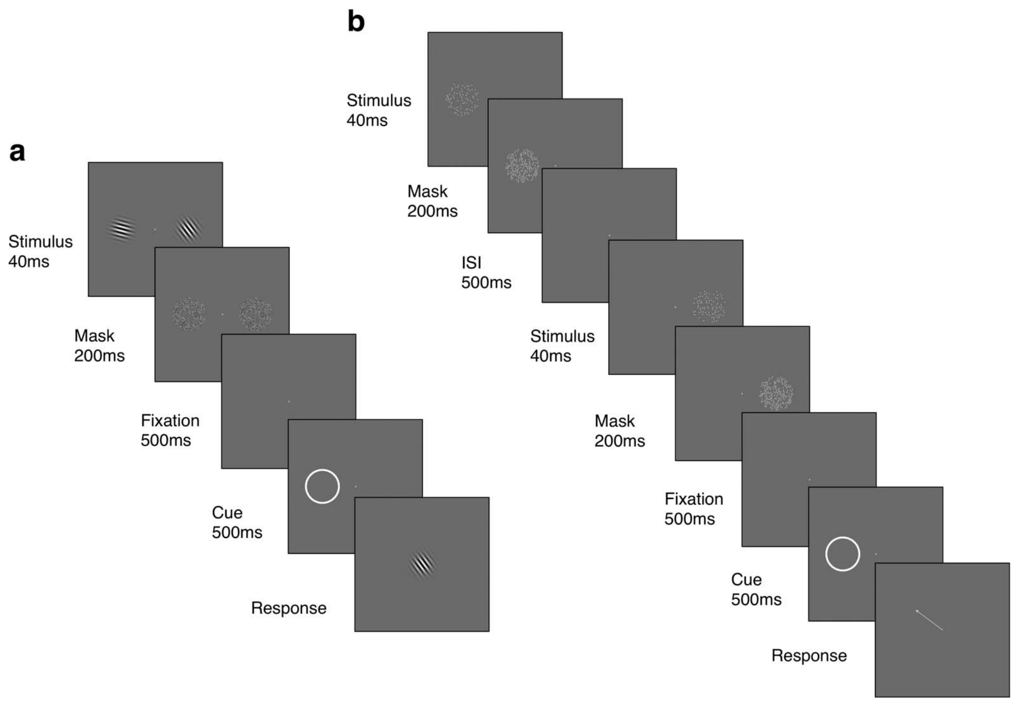

Figure 1. Examples of the presentation sequence used in Experiment 1. An example of simultaneous presentation of orientation items (left) and sequential presentation of motion direction items (right).

consolidates the cued item into VWM, resulting in a normal distribution of offsets with a mean (μ) and standard deviation (σ); the other where they fail to consolidate the item and must guess (g), resulting in a rectangular or even distribution. Thus, in order to examine whether there is a precision cost associated with parallel consolidation, a mixture model must be fit to the offset data to isolate these sources of variation. A model was fit to individual offset data within each feature condition using a standard maximum-likelihood method. Data analysis was performed using the MemToolbox (Suchow, Brady, Fougnie, & Alvarez, 2013).

Results

Descriptive statistics

The raw offset data for orientation and motion direction was analyzed separately for bias (mean) and variability (variance) using a one-way, repeated measures analysis of variance (ANOVA). One observer (in the orientation condition) reported being unable to perform the task even in the sequential condition; the data reflected this (flat distribution of offsets) and was omitted from analysis. The mean of the offset data was equivalent between presentation conditions for both orientation and direction, $F(1, 10) = 0.1$ and $F(1, 11) = 0.3$, respectively, and one-sample t tests revealed that none of the means differed significantly from zero (all $p\text{s} > 0.17$) (Figure 2a). In contrast, variance differed considerably between conditions (Figure 2b); however, because it was not normally distributed, the variance was transformed by taking the logarithm prior to analysis. The log variance differed significantly between conditions for both features, $F(1, 10) = 13.3$ $p < 0.01$ (orientation) and $F(1, 11) = 9.0$ $p < 0.05$ (direction); see Figure 2c. These results show that the offset between target and response orientation/direction was more variable when items were presented simultaneously than sequentially. This pattern of results was highly consistent across individual participants (Figures 2e and 2d).

Model fit

A mixed model was fit to individual observer's data and statistical analysis was performed on model parameters to evaluate how the guess rate (g) and standard deviation (σ) of responses varied between sequential and simultaneous presentation of items. The Bayesian information criterion (BIC) was used to compare the fit of three types of mixed models: standard mixture model, variable precision model, and swap model. Whereas a standard mixture model assumes precision remains constant, a variable precision model assumes precision is normally distributed and calculates its mean and standard deviation (Fougnie, Suchow, & Alvarez, 2012). In addition to partitioning sources of variance into guess rate and precision, like the former two models, a swap model

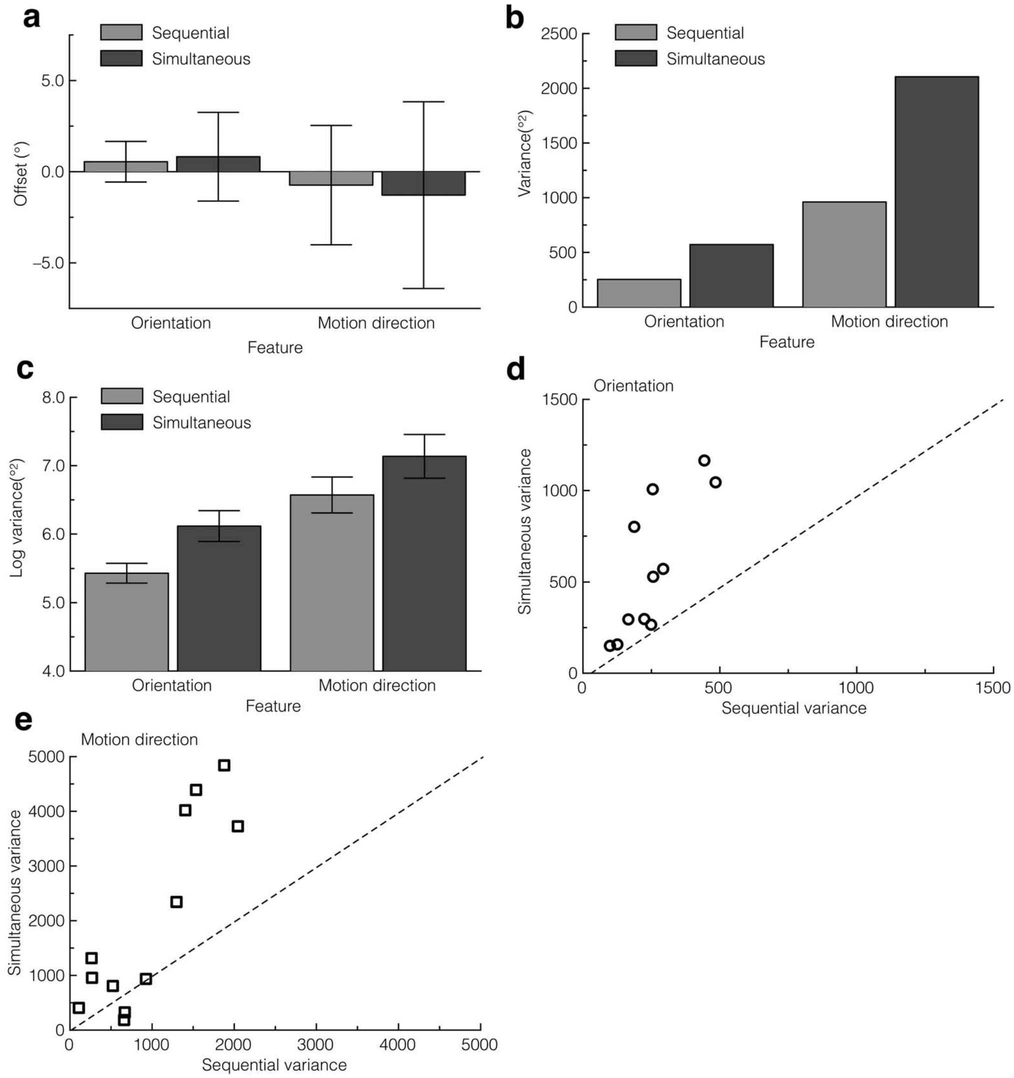

Figure 2. Experiment 1 raw offset data. The (a) grand-average offset between target and response orientation/direction, (b) average variance in offset, and (c) average log variance in offset. Data from sequential and simultaneous conditions are represented by light and dark grey bars, respectively, in (a), (b), and (c); error bars in (a) and (c) represent ±1 *SEM*. Scatterplots (d) and (e) show individual observer variance in the simultaneous condition as a function of variance in the sequential condition; circles and squares represent observers in the orientation and motion direction conditions, respectively. Points above the dotted line indicate higher variance in the simultaneous condition relative to the sequential condition.

isolates a third potential source of variation: Responses made based on the orientation/direction of the nontarget item (Bays, Catalao, & Husain, 2009). A standard mixture model was used as it was found to fit individual observer data better than the variable precision and swap models: Standard mixture model BIC scores were lowest for over 90% of data sets (Supplementary material). In the VWM storage literature, studies conducting more systematic model comparisons have tended to reject the mixture model in favor of other models, e.g., swap and variable precision (Fougnie et al., 2012; Van den Berg, Awh, & Ma, 2014; Sims, 2015). Here, the rejection of these models in favor of the mixture model may reflect differences in the nature of the cognitive process being investigated: VWM consolidation, as opposed to storage.

Model fit was evaluated using Kolmogorov-Smirnov (K-S) tests, revealing that the standard mixture model fit the data well (all ps > 0.3). This model decomposes data into a mixture of parameters that are characterized by either a uniform or von Mises distribution of errors (Zhang & Luck, 2008). Note that although we found a mixture model fit the data best, we cannot rule out that a proportion of the responses categorized as guesses were actually a result of spatial binding errors (swapping), simply that this proportion was not sufficiently large enough to tip the balance in favor of the swap model during model comparison. Both the guess rate (g) and standard deviation (σ) parameters

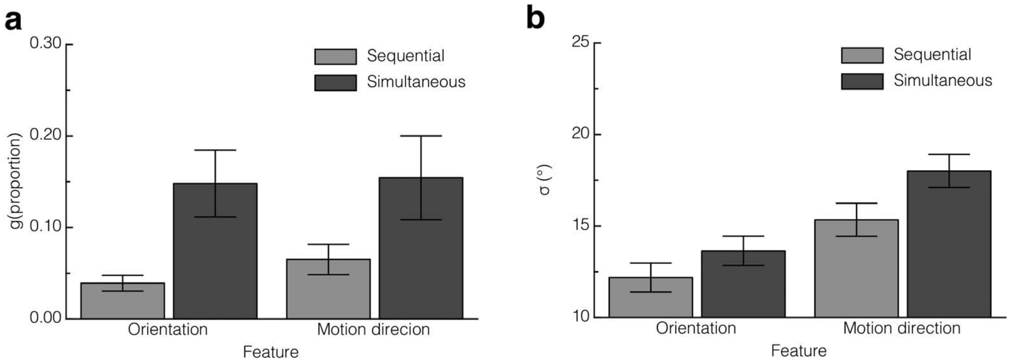

Figure 3. Experiment 1 model parameter analysis. The average (a) guess rate (g) and (b) standard deviation (σ) model parameters of observers in sequential/simultaneous conditions for orientation and motion direction. All error bars indicate ± 1 *SEM*.

significantly increased from sequential to simultaneous presentation for orientation, $t(10) = 3.4$ $p < 0.01$ and $t(10) = 2.3$ $p < 0.05$, and direction, $t(11) = 2.7$ $p < 0.05$ and $t(11) = 3.6$ $p < 0.01$, respectively (Figure 3). Note that the standard deviation (σ) parameter is an inverse measure of precision, i.e., higher values indicate poorer precision.

Given the average retention interval, i.e., the duration between each item exposure and the response interval, was less in the simultaneous (1200 ms) than the sequential condition (1550 ms), these results must be due to processes impacted at the consolidation stage, not during storage. This hypothesis is further evidenced by the similarity across all parameters between models fit to the trials where the cued item was in the first compared to the second interval, in the sequential condition, for both features (all ps > 0.05).

Discussion

The increased rate of guessing found in the simultaneous conditions across both features suggests that observers may not have been capable of parallel consolidation of these features. A strictly serial condition strategy predicts a guess rate of at least 50% in the simultaneous condition, reflecting failure to consolidate more than one item on each trial. However, the guess rate in these conditions is considerably less than 50%. Two possible explanations could account for this result. Either observers are capable of performing parallel consolidation of these features, but incur an increased likelihood of consolidation failure as a result, or observers are not able to consolidate items in parallel but the exposure duration employed allowed them to serially consolidate a second item on a number of trials. The exposure duration (40 ms) used in Experiment 1was less than that used in a previous study in which the authors claimed the duration (150 ms) was sufficiently short to prevent serial consolidation of more than one item (Liu & Becker, 2013). Thus, it seems surprising that observers would have been capable of serially consolidating items in the simultaneous conditions here. However, given that the average guess rate in the sequential conditions is around 5%, it is possible that the difference in guess rate between the sequential and simultaneous conditions may have been underestimated due to a ceiling effect in the sequential conditions.

The increased standard deviation found in the simultaneous conditions is compelling evidence that observers were employing parallel consolidation, and as a result, items were encoded at a lower precision. It is likely that this is due to spreading of cognitive resources employed during consolidation. However, another potential mechanism for this precision loss concerns the spatial attention that can be employed to enhance processing of the items during encoding. That is, given that the order and location of item presentation was consistent throughout the experiment, observers may have been making covert attentional shifts to the locations of the items in the sequential condition. In contrast, in the simultaneous condition attention would be spread across the two locations, resulting in less effective facilitation of spatial attention (Castiello & Umiltà, 1992; Eriksen & Yeh, 1985). This facilitation would explain why in previous studies, where the location of presented items was randomized, no difference in precision was found (Liu & Becker, 2013; Miller et al., 2014).

The current results provide partial evidence that both orientation and motion direction can be consolidated in parallel; that is, the difference in guess rate is less than would be predicted by a serial consolidation strategy and, more importantly, modulation of precision between sequential and simultaneous conditions was found, indicating loss of precision resulting from parallel consolidation. However, the difference in guess rate between sequential and simultaneous conditions may have been underestimated due to an overly long exposure duration, and the modulation of precision may have been due to facilitation of covert attentional shifts in the sequential conditions. Experiment 2 was run to investigate these possibilities.

Experiment 2

Although the modulation in precision found between presentation conditions in Experiment 1 suggests that observers were performing parallel consolidation of orientation and motion direction, this may have been a result of covert attentional shifts in the sequential condition. To examine this possibility we compared precision between fixed and random sequential presentation. If the difference in precision is due to covert attention, we should observe better precision in the fixed sequential condition, compared to the random sequential condition. However, if it is due to parallel consolidation, we would expect that precision in the simultaneous condition will be less than in both sequential conditions.

Furthermore, the magnitude of the difference in guess rate between the sequential and simultaneous conditions in Experiment 1 may have been underestimated. That is, there may have been a ceiling effect in the sequential condition due to an overly long exposure duration, which may also have resulted in observers employing serial consolidation in the simultaneous condition. Here we investigate this possibility by tailoring the exposure duration of the stimuli to each individual, in order to bring performance in the sequential condition to threshold and ensure serial consolidation cannot be used in the simultaneous condition.

Method

Observers

Twenty-four observers participated in the study (mean age, 22). All had normal or corrected-to-normal acuity and gave informed written consent to participate in the study. All observers were compensated $20 for participation.

Stimuli and procedure

The stimuli and procedure were the same as that used in Experiment 1, with a few notable exceptions. To examine whether the difference in precision found between sequential and simultaneous conditions in Experiment 1 resulted from covert shifts of attention in the sequential condition, here we ran two sequential presentation conditions: one with fixed presentation order (replicating Experiment 1) and one with randomized presentation order.

In order to calibrate the exposure duration of the stimuli such that performance would be closer to threshold in the sequential condition, and thus examine the possibility that the difference in guess rate between sequential and simultaneous conditions in Experiment 1 was underestimated, a threshold exposure duration was determined for each observer before running the main experiment. The threshold exposure duration stimuli and procedure were the same as those used in the sequential condition of the main experiment, except that now an adaptive staircase procedure was employed using software from the Palamedes Toolbox (Prins & Kingdom, 2009), varying the exposure duration of the items. The staircase uses a "psi-marginal" adaptive method, based on Kontsevich and Tyler's (1999) psi-method, which allows any of the four parameters of the psychometric function to be treated as a parameter of primary interest, a "nuisance" parameter, or a fixed parameter (Prins, 2013). Each staircase consisted of 50 trials and responses were considered correct if they were within 30° of the target orientation/direction (i.e., $\sim \pm 2$ standard deviation found in Experiment 1). This resulted in a chance level of 0.33 and 0.16 for orientation and direction, respectively; thus, the threshold performance levels used were 0.66 and 0.58, respectively.

Here, as in Experiment 1, observers were randomly split into two groups: Half were run in the orientation condition and the other half in the direction condition. Initially, observers' exposure duration threshold was determined using the previously reported staircase procedure. Following this, each observer ran six blocks of each presentation condition (i.e., fixed/random sequential and simultaneous) randomly interleaved within a mega block. Each block consisted of 50 trials, totalling 900 trials and an approximate testing duration of one hour per observer.

Results

Threshold exposure duration

The average threshold duration was 43.3 ms (range, 16–88 ms; *SD*, 20.6 ms) for orientation and 89.3 ms (range, 40–160 ms; *SD*, 40.1 ms) for direction. This is similar to the exposure duration for these features found in previous studies (Becker et al., 2013; Rideaux et al., 2015).

Descriptive statistics

The raw offset data for orientation and motion direction was analyzed separately for bias and variance using a one-way repeated measures ANOVA. Because it was not normally distributed, variance was transformed by taking the logarithm prior to analysis. No

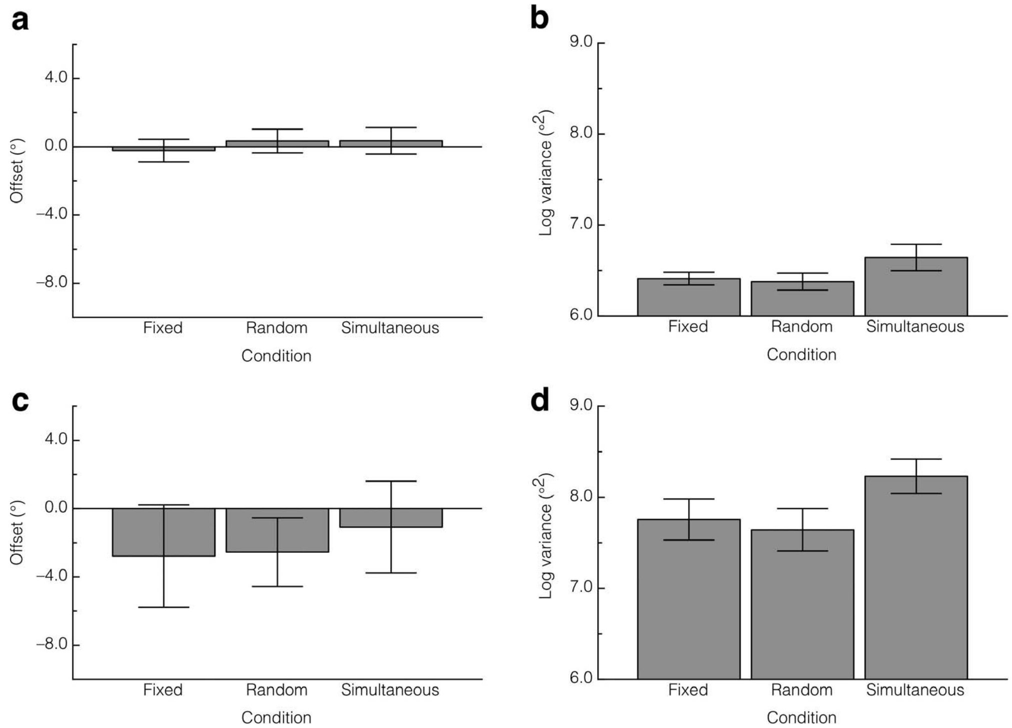

Figure 4. Experiment 2 raw offset data. The grand-average offset between target and response and average log variance in offset for orientation (a) and (b) and direction (c) and (d), respectively. All error bars represent ±1 *SEM*.

main effects of mean or log variance were found for orientation between the three presentation conditions, $F(2, 11) = 0.4$ $p = 0.69$ and $F(2, 11) = 2.9$ $p = 0.08$, respectively (Figures 4a and 4b). Whereas no main effect of mean was found for direction, $F(2, 11) = 0.4$ $p = 0.66$, a significant main effect of log variance was found, $F(2, 11) = 28.3$ $p < 0.001$ (Figures 4c and 4d). Thus, whereas the preliminary results for direction mirror those found in Experiment 1, those for orientation suggest that the offset between target and response was similarly variable between sequential and simultaneous presentation conditions, and between fixed and random sequential presentation conditions.

Model fit

To evaluate the differences in guess rate and standard deviation between presentation conditions, a standard mixture model was fit to individual's offset data, and a one-way repeated measures ANOVA was run on each parameter. Consistent with Experiment 1, a mixture model was used as it was found to fit individual observer data better than the variable precision and swap models: Standard mixture model BIC scores were lowest for over 95% of data sets (Supplementary material), and overall the models fit the data well (95% of ps > 0.05, assessed using K-S tests).

For orientation, no main effect of standard deviation was found, $F(2, 11) = 1.3$ $p = 0.28$, and although precision is poorest in the sequential random and simultaneous conditions, none of the differences between conditions were significance (all ps > 0.15, assessed using paired t tests) (Figure 5a). In contrast, a main effect of standard deviation was found for motion direction, $F(2, 11) = 17.6$ $p < 0.001$, with paired t tests revealing significant differences between fixed/random sequential and simultaneous conditions, $t(11) = 5.2$ $p < 0.001$ and $t(11) = 4.7$ $p = 0.001$, respectively, but no difference between sequential conditions, $t(11) = 0.1$ $p = 0.92$, (Figure 5c) . Thus, these results show that the modulation in precision found in Experiment 1 (at least for motion direction) was not a result of covert attentional shifts, but likely due to spreading of other cognitive resources engaged during consolidation.

Significant main effects of guess rate were found for both orientation, $F(2, 11) = 4.2$ $p = 0.02$, and motion direction, $F(2, 11) = 12.3$ $p < 0.001$ (Figure 5b and 5d). Paired t test revealed a similar pattern of results for both features—no significant differences between fixed and random sequential conditions (all ps > 0.15), and differences between sequential and simultaneous conditions were all significant (all ps < 0.05) with the exception of that between random sequential and simultaneous conditions for orientation, $t(11) = 2.1$ $p = 0.06$.

Comparative analysis of parameters derived from models fit to the trials where the cued item was in the first or second interval (in the sequential condition) yielded similar results to those found in Experiment 1.

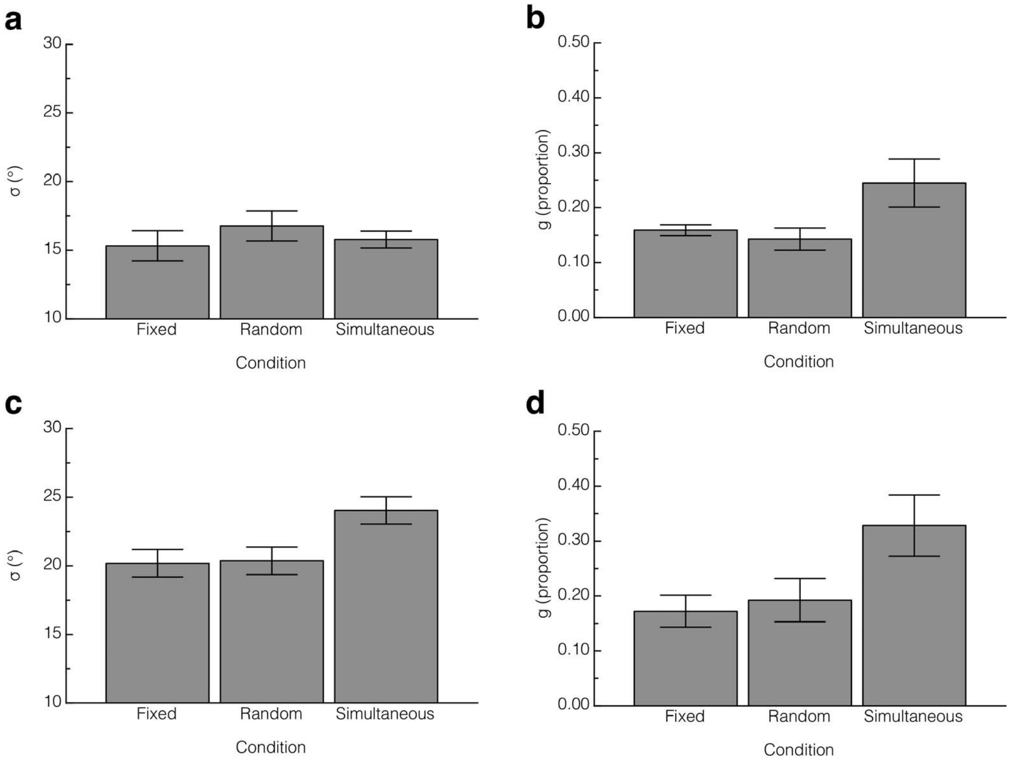

Figure 5. Experiment 2 model parameter analysis. The average standard deviation (σ) and guess rate (g) model parameters of observers in (fixed and random) sequential and simultaneous presentation conditions for orientation (a) and (b) and motion direction (c) and (d), respectively. All error bars indicate ± 1 *SEM*.

That is, no difference between parameters (all *p*s $>$ 0.05), with the exception of the standard deviation of orientation in the random sequential condition which increased significantly when the cued item was presented in the second interval, $t(11) = 2.6$ $p = 0.02$. In this condition, the spatial location of the first item could not be anticipated, but the location of the second item could, as there were only two possible locations and items were not presented in the same location within a trial. This finding suggests that the capacity to anticipate the location where the orientation item was to be presented worsened the precision at which it was encoded. Alternatively, this finding could suggest that having an existing item stored in VWM reduces the precision of subsequently stored items; however, as this was not replicated in the fixed condition or the previous experiment, this theory seems less likely.

Discussion

In Experiment 1, although the average guess rate in the simultaneous condition was significantly below the most conservative estimate predicted by a serial consolidation strategy (50%), it was also significantly higher than the guess rate in the sequential condition. Furthermore, this differential may have been underestimated due to a ceiling effect in the sequential condition, in which the average guess rate was around 5%. That is, whereas performance indicated observers were capable of parallel consolidation in the simultaneous condition, performance in the sequential condition suggested that the exposure duration may have been sufficient to consolidate more than one item serially. In the current experiment, given that the average guess rate in the sequential conditions is around 15%–20%, performance in these conditions cannot reflect a ceiling effect.

For motion direction, the modulation of precision was replicated in Experiment 2. Furthermore, the similarity between precision in the fixed and random sequential conditions demonstrates that this difference is not due to covert attentional shifts. Thus, this is convincing evidence that motion direction can be consolidated in parallel and that as a result, items are encoded at a reduced precision.

In contrast, for orientation, as the result of tailoring the exposure duration to individual observers, precision was not modulated here, as in Experiment 1. Given that no difference in precision was found here between fixed and random sequential presentation, it is unlikely that modulation of precision in Experiment 1 was a result of covert attentional shifts. Rather, it is possible that in Experiment 1, the fixed duration employed allowed

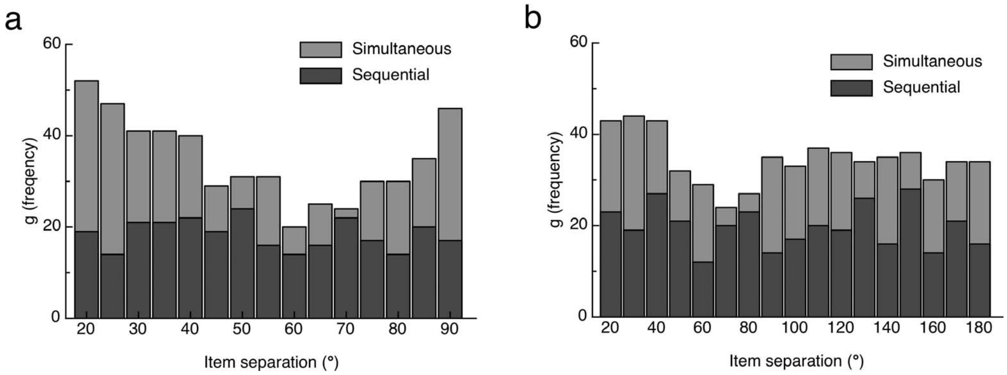

Figure 6. Frequency of guess responses as a function of the angular separation between presented items in the (a) orientation and (b) motion direction conditions of Experiment 1.

(some) observers to serially consolidate two items in the simultaneous presentation consolidation; however, items were consolidated in a shorter duration, resulting in lower precision encoding. This, in addition to the increased guess rate in the simultaneous condition, could indicate that observers are limited to serial consolidation of orientation, consistent with previous research (Liu & Becker, 2013). However, this strategy would predict a guess rate of at least 50%, which is considerably more than what we observed (~25%), showing that on a number of trials observers were capable of consolidating both items in the simultaneous condition.

One possibility is that certain combinations of orientations, e.g., horizontal and vertical, can be consolidated as one item due to their activating higher level structures, e.g., a cross. Indeed, a number of studies have found evidence supporting summary statistics or hierarchical representations in VWM (Brady & Alvarez, 2011; Brady & Tenenbaum, 2013; Orhan & Jacobs, 2013; Orhan, Sims, Jacobs, & Knill, 2014). To evaluate this possibility, we included the midpoint of orientation/direction items within each trial as a possible swap model "distractor" and compared the fit of this new "averaging model" with the standard mixture model. No evidence was found that the averaging model could explain the data better than the standard mixture model: Mixture model BIC scores were lowest for over 99% of data sets in Experiments 1 and 2 (Supplementary material), thus it seems unlikely that these results can be accounted for by hierarchical representations/ summary statistics.

Alternatively, the results may indicate that orientation, like motion direction and color, can be consolidated in parallel, but suffers an increased likelihood of consolidation failure as a result of simultaneous presentation. Indeed, it is interesting that for both features there was a significantly higher likelihood of consolidation failure when items were presented simultaneously. However, as previously mentioned, there is convincing evidence to suggest that the guess rate would be higher in the simultaneous condition (Ihssen et al., 2010; Scalf & Beck, 2010). That is, studies show that presenting items simultaneously results in increased likelihood of consolidation failure due to competitive interference between representations. This interference is known to be influenced by the similarity of items (Shapiro & Miller, 2011).

In order to examine whether this interference could account for the difference in guess rate between sequential and simultaneous conditions, we plotted the known guess responses across all observers in Experiment 1 as a function of the angular difference between items on corresponding trials. Responses were considered guesses if they fell more than two standard deviations (derived from the model) away from the target orientation/direction. The results of this analysis are presented in Figure 6. There appears to be no relationship between item similarity and likelihood of consolidation failure in the sequential conditions, indicated by a flat distribution. In contrast, it appears that there is a relationship between these factors in the simultaneous conditions such that items of greater similarity are more likely to result in interference between 20°–70°, with this relationship reversing with separation greater than 70° (plateauing after 115° for direction). The evidence of interference within the simultaneous condition, but not the sequential condition, is consistent with biased competition models of VWM/attention and would explain the difference in guess rate between these conditions within a framework of parallel consolidation.

Differentiating between serial and parallel models is often challenging; however, modulation of precision is compelling evidence for the latter. Whereas we found no difference in precision between conditions for orientation, the results for motion direction clearly indicate that observers were capable of parallel consolidation and, as a result, items were encoded at a reduced precision.

General discussion

The main findings were that motion direction can be consolidated in parallel and that there is a twofold cost: reduced precision encoding and an increase in consolidation failure. The evidence found for orientation was less conclusive and could plausibly be explained by either a serial or a parallel account. The reduction in precision observed for motion direction is likely due to spreading of cognitive processes associated with parallel consolidation. For instance, the implicit goals of observers may have differed between conditions (Sims, 2015), i.e., devaluing precision in the simultaneous condition in order to achieve parallel consolidation. As evidenced by the post hoc analysis of guess responses and item similarity, the increase in consolidation failrate may be due to interference between items presented simultaneously, as opposed to sequentially.

In our previous study, we suggested that a reduction in the precision of items consolidated in parallel may account for the difference in performance observed between sequential and simultaneous conditions. The results of the current study confirm this interpretation for motion direction. The previous results were also suggestive that both orientation and motion direction can be consolidated in parallel, with stronger evidence for direction than orientation. Although here we have found compelling evidence for parallel consolidation of motion, once again the results for orientation are less conclusive.

Previous research indicated that whereas color can be consolidated in parallel (at no cost), orientation is limited to serial consolidation (Liu & Becker, 2013). An all-or-none "unlimited parallel" model of consolidation was proposed to account for these results, where it was claimed that the information bandwidth of color was small enough that two items could pass through simultaneously, while the bandwidth of orientation was too large to accomplish this (Miller et al., 2014). As this model does not predict any cost of parallel consolidation, it cannot explain the current findings: that is, the reduction in precision observed when motion direction items are consolidated in parallel. The information bandwidth model is resonant of current discrete models of VWM storage, characterized by precision invariant storage in a discrete number of "slots." However, the current findings are more parsimoniously explained by consolidation that draws upon a continuous resource, which can be allocated among a number of items, with a relationship between resource allocation and consolidation precision.

A possible explanation for the apparent discrepancy between the cost of parallel consolidation for color and motion direction is that color is processed more categorically than motion direction and thus less susceptible to precision decay. There is some evidence for this from event-related potential (ERP) studies where the pattern of results observed when the contralateral delay activity (CDA), a physiological indicator of both the number and precision of items stored in VWM, is measured while storing either orientation or color in VWM. Whereas the pattern of results for orientation reflect a continuous resource model of VWM storage (Gao, Yin, Xu, Shui, & Shen, 2011), results using color reflect a discrete model (Ikkai, McCollough, & Vogel, 2010; Luria, Sessa, Gotler, Jolicoeur, & Dell'Acqua, 2010; Ye, Zhang, Liu, Li, & Liu, 2014). However, it is difficult to make direct comparisons, as this technique has not yet been used to investigate VWM storage of motion direction. Furthermore, it is important to note that numerous behavioral studies investigating VWM storage of color, orientation, and motion stimuli have reported a pattern of results consistent with a resource model (Bays & Husain, 2008; Bays et al., 2009; van den Berg, Shin, Chou, George, & Ma, 2012; Zokaei, Gorgoraptis, Bahrami, Bays & Husain, 2011).

We also found that, up to around 70°, similar items were more susceptible to consolidation failure; this is consistent with our previous study where we found that reducing the separation between motion direction items (from 90° to 45°) resulted in a differential between sequential and simultaneous conditions. This may also explain why here we found increased consolidation failure for orientation and motion direction, which have relatively small perceptual spaces (Clifford, 2002; Foster & Ward, 1991; Webster, De Valois, & Switkes, 1990), and why no difference was found for color (Miller et al., 2014), which has a relatively large perceptual space (Nagy & Sanchez, 1990; Witzel & Gegenfurtner, 2013). Indeed, the minimum separation between colors presented by Miller et al. (2014) was relatively large, i.e., 45° on the color wheel; perhaps reducing this would result in the same increase in consolidation failure observed here.

In summary, the current findings are consistent with our previous study indicating that motion direction can be consolidated into VWM in parallel (Rideaux et al., 2015). However, we extend this by demonstrating that, unlike color, there is a twofold cost associated with parallel consolidation of motion direction: The precision at which items are encoded is reduced and the likelihood of consolidation failure is increased. Evidence is also found suggesting that parallel consolidation of orientation may be possible, but is not conclusive. These findings emphasize that parallel consolidation is not unique to color, and suggest that part of the cost of parallel consolidation may be mediated by the size of the perceptual space of these features.

Keywords: motion direction, orientation, parallel consolidation, visual working memory, biased competition model

Acknowledgments

This work was supported by an Australian Postgraduate Award to R. R. and an Australian Research Council Grant (DP110104553) to M. E.

Commercial relationships: none
Corresponding author: Reuben Rideaux.
Email: reuben.rideaux@anu.edu.au.
Address: Research School of Psychology, The Australian National University, Canberra, Australia.

References

Albright, T. D., Desimone, R., & Gross, C. G. (1984). Columnar organization of directionally selective cells in visual area MT of the macaque. *Journal of Neurophysiology*, 51(1), 16–31. Retrieved from http://jn.physiology.org/content/51/1/16.full-text.pdf+html

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10):7, 1–11, doi:10.1167/9.10.7. [PubMed] [Article]

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*(5890), 851–854, doi:10.1126/science.1158023.

Beck, D. M., & Kastner, S. (2007). Stimulus similarity modulates competitive interactions in human visual cortex. *Journal of Vision*, *7*(2):19, 1–12, doi:10.1167/7.2.19. [PubMed] [Article]

Becker, M. W., Miller, J. R., & Liu, T. (2013). A severe capacity limit in the consolidation of orientation information into visual short-term memory. *Attention, Perception & Psychophysics*, *75*(3), 415–425, doi:10.3758/s13414-012-0410-0.

Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory ensemble statistics bias memory for individual items. *Psychological Science*, *22*(3), 384–392, doi:10.1177/0956797610397956.

Brady, T. F., & Tenenbaum, J. B. (2013). A probabilistic model of visual working memory: Incorporating higher order regularities into working memory capacity estimates. *Psychological review*, *120*(1), 85–109, doi:10.1037/a0030779.

Brainard, D. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. Retrieved from http://bbs.bioguider.com/images/upfile/2006-4/200641014348.pdf

Castiello, U., & Umiltà, C. (1992). Splitting focal attention. *Journal of Experimental Psychology. Human Perception and Performance*, *18*, 837–848, doi:10.1037/0096-1523.18.3.837.

Clifford, C. W. (2002). Perceptual adaptation: motion parallels orientation. *Trends in Cognitive Sciences*, *6*(3), 136–143, doi:10.1016/S1364-6613(00)01856-8.

Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and Brain Sciences*, *24*(1), 154–176. Retrieved from http://journals.cambridge.org/abstract_S0140525X0161392X

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, *19*(1), 51–57, doi:10.1177/0963721409359277.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18(1), 193–222, doi:10.1146/annurev.ne.18.030195.001205.

Edwards, M., & Greenwood, J. A. (2005). The perception of motion transparency: A signal-to-noise limit. *Vision Research*, *45*(14), 1877–1884, doi:10.1016/j.visres.2005.01.026.

Edwards, M., & Rideaux, R. (2013). How many motion signals can be simultaneously perceived? *Vision Research*, *76*, 11–16, doi:10.1016/j.visres.2012.10.004.

Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology Human Perception and Performance*, *11*, 583–597, doi:10.1037/0096-1523.11.5.583.

Foster, D. H., & Ward, P. A. (1991). Asymmetries in oriented-line detection indicate two orthogonal filters in early vision. *Proceedings of the Royal Society of London B: Biological Sciences*, *243*(1306), 75–81, doi:10.1098/rspb.1991.0013.

Fougnie, D., Suchow, J. W., & Alvarez, G. A. (2012). Variability in the quality of visual working memory. *Nature Communications*, *3*, 1229, doi:10.1038/ncomms2237.

Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: competitive content maps for attention and memory. *Trends in Cognitive Sciences*, *17*(3), 134–141, doi:10.1016/j.tics.2013.01.010.

Gao, Z., Yin, J., Xu, H., Shui, R., & Shen, M. (2011). Tracking object number or information load in visual working memory: Revisiting the cognitive implication of contralateral delay activity. *Biolog-*

ical Psychology, *87*(2), 296–302, doi:10.1016/j.biopsycho.2011.03.013.

Huang, L., Treisman, A., & Pashler, H. (2007). Characterizing the limits of human visual awareness. *Science*, *317*(5839), 823–825, doi:10.1126/science.1143515.

Ihssen, N., Linden, D. E., & Shapiro, K. L. (2010). Improving visual short-term memory by sequencing the stimulus array. *Psychonomic Bulletin & Review*, *17*(5), 680–686, doi:10.3758/PBR.17.5.680.

Ikkai, A., McCollough, A. W., & Vogel, E. K. (2010). Contralateral delay activity provides a neural measure of the number of representations in visual working memory. *Journal of Neurophysiology*, *103*(4), 1963–1968, doi:10.1152/jn.00978.2009.

Kastner, S., & Ungerleider, L. G. (2001). The neural basis of biased competition in human visual cortex. *Neuropsychologia*, *39*(12), 1263–1276, doi:10.1016/S0028-3932(01)00116-6.

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, *39*(16), 2729–2737, doi:10.1016/S0042-6989(98)00285-5.

Liu, T., & Becker, M. W. (2013). Serial consolidation of orientation information into visual short-term memory. *Psychological Science*, *24*(6), 1044–1050, doi:10.1177/0956797612464381.

Luria, R., Sessa, P., Gotler, A., Jolicœur, P., & Dell'Acqua, R. (2010). Visual short-term memory capacity for simple and complex objects. *Journal of Cognitive Neuroscience*, *22*(3), 496–512, doi:10.1162/jocn.2009.21214.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356, doi:10.1038/nn.3655.

Mance, I., Becker, M. W., & Liu, T. (2012). Parallel consolidation of simple features into visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 429–438, doi:10.1037/a0023925.

Miller, J. R., Becker, M. W., & Liu, T. (2014). The bandwidth of consolidation into visual short-term memory depends on the visual feature. *Visual Cognition*, *22*(7), 920–947. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/13506285.2014.936923#.VMb18XayxoU

Nagy, A. L., & Sanchez, R. R. (1990). Critical color differences determined with a visual search task. *Journal of the Optical Society of America A*, *7*(7), 1209–1217, doi:10.1364/JOSAA.7.001209.

Orhan, A. E., & Jacobs, R. A. (2013). A probabilistic clustering theory of the organization of visual short-term memory. *Psychological Review*, *120*(2), 297–328, doi:10.1037/a0031541.

Orhan, A. E., Sims, C. R., Jacobs, R. A., & Knill, D. C. (2014). The adaptive nature of visual working memory. *Current Directions in Psychological Science*, *23*(3), 164–170, doi:10.1177/0963721414529144.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442, doi:10.1163/156856897X00366.

Prins, N. (2013). The psi-marginal adaptive method: How to give nuisance parameters the attention they deserve (no more, no less). *Journal of Vision*, *13*(7): 3, 1–17, doi:10.1167/13.7.3. [PubMed] [Article]

Prins, N., & Kingdom, F. A. A. (2009). Palamedes: Matlab routines for analyzing psychophysical data. Retrieved from http://www.palamedestoolbox.org

Qian, N., Andersen, R., & Adelson, E. (1994). Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics. *Journal of Neuroscience*, *14*(12), 7357–7366. Retrieved from http://www.jneurosci.org/content/14/12/7357.short [PubMed] [Article]

Rideaux, R., Apthorp, D., & Edwards, M. (2015). Evidence for parallel consolidation of motion direction and orientation into visual short-term memory. *Journal of Vision*, *15*(2):17, 1–12, doi:10.1167/15.2.17.

Rideaux, R., & Edwards, M. (2014). Information extraction during simultaneous motion processing. *Vision Research*, *95*, 1–10, doi:10.1016/j.visres.2013.11.007.

Scalf, P. E., & Beck, D. M. (2010). Competition in visual cortex impedes attention to multiple items. *The Journal of Neuroscience*, *30*(1), 161–169, doi:10.1523/JNEUROSCI.4207-09.2010.

Shapiro, K. L., & Miller, C. E. (2011). The role of biased competition in visual short-term memory. *Neuropsychologia*, *49*(6), 1506–1517, doi:10.1016/j.neuropsychologia.2011.02.017.

Sims, C. R. (2015). The cost of misremembering: Inferring the loss function in visual working memory. *Journal of Vision*, *15*(3):2, 1–27, doi:10.1167/15.3.2. [PubMed] [Article]

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*, 1–29, doi:10.1037/h0093759.

Sperling, G. (1963). A model for visual memory tasks.

Human Factors: The Journal of the Human Factors and Ergonomics Society, *5*(1), 19–31, doi:10.1177/001872086300500103.

Suchow, J. W., Brady, T. F., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision*, *13*(10):9, 1–8, doi:10.1167/13.10.9. [PubMed] [Article]

Umemoto, A., Drew, T., Ester, E. F., & Awh, E. (2010). A bilateral advantage for storage in visual working memory. *Cognition*, *117*(1), 69–79, doi:10.1016/j.cognition.2010.07.001.

van den Berg, R., Awh, E., & Ma, W. J. (2014). Conceptualizing and testing working memory models in a three-dimensional model space. *Psychological Review*, *121*(1), 124–149, doi:10.1037/a0035234.

van den Berg, R., Shin, H., Chou, W. C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences, USA, 109*(22), 8780–8785, doi:10.1073/pnas.1117465109.

Webster, M. A., De Valois, K. K., & Switkes, E. (1990). Orientation and spatial-frequency discrimination for luminance and chromatic gratings. *Journal of the Optical Society of America A*, *7*(6), 1034–1049, doi:10.1364/JOSAA.7.001034.

Witzel, C., & Gegenfurtner, K. R. (2013). Categorical sensitivity to color differences. *Journal of Vision*, *13*(7):1, 1–33, doi:10.1167/13.7.1. [PubMed] [Article]

Ye, C., Zhang, L., Liu, T., Li, H., & Liu, Q. (2014). Visual working memory capacity for color is independent of representation resolution. *PloS one*, *9*(3), e91681, doi:10.1371/journal.pone.0091681.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235, doi:10.1038/nature06860.

Zokaei, N., Gorgoraptis, N., Bahrami, B., Bays, P. M., & Husain, M. (2011). Precision of working memory for visual motion sequences and transparent motion surfaces. *Journal of Vision*, *11*(14):2, 1–18, doi:10.1167/11.14.2. [PubMed] [Article]

Supplementary material

All Bayesian information criterion values, evaluated for the purpose of model comparisons, are provided as Supplementary material.